SHORT COMMUNICATION

ANIMAL GENETICS WILEY

Inbreeding depression and durability in the North American Thoroughbred horse

Abstract

The proportion of the genome containing runs of homozygosity (ROH) affects production traits in livestock populations. In European and Australasian Thoroughbreds inbreeding, quantified using ROH (F_{ROH}), is associated with the probability of ever racing. Here, we measured F_{ROH} using 333 K SNP genotypes from 768 Thoroughbred horses born in North America to evaluate the effect of inbreeding on racing traits in that region. Among North American horses, F_{ROH} was not associated ($p = 0.518$) with the probability of ever racing but was significantly associated with the number of race starts ($p = 0.002$). Among raced horses, those with a 10% higher F_{ROH} than the mean inbreeding coefficient were predicted to have 3.5 fewer race starts compared to horses with a mean inbreeding coefficient. Considering the trend of increasing inbreeding and a decline in the average number of race starts per runner in North America, mitigating inbreeding in the population could positively influence racing durability.

The Thoroughbred horse has low genetic diversity relative to most other horse breeds, with a small effective population size (Orlando & Librado, 2019; Petersen et al., 2013) and a trend of increasing inbreeding (McGivney et al., 2020). Until recently, the effects of inbreeding in the population were unknown. In the Australasian Thoroughbred, pedigree-based estimates of inbreeding are associated with a strong negative effect on racing performance (Todd et al., 2018). In European and Australasian Thoroughbreds inbreeding, quantified based on runs of homozygosity (ROH) determined from genome-wide SNP data, is associated with a lower probability of ever racing (Hill et al., 2022).

Although the population is globally distributed, the Thoroughbred is largely genetically homogeneous with considerable overlap in variation among regions (McGivney et al., 2020). For each geographical region, the prominent sire lines are the principal drivers of genetic diversity (McGivney et al., 2020). Nonetheless, region-specific selection for traits favourable to the different racing ecosystems (Han et al., 2020) and variation in breeding practices, may impact on differences in population genetics across geographic regions (Lampi et al., 2020).

Therefore, we examined the effects of inbreeding in the North American Thoroughbred population and compared this with results obtained for European and Australasian horses. We quantified inbreeding based on ROH (F_{ROH}) using 333 K SNP genotypes from $n = 768$ horses ($n = 207$ males, $n = 561$ females) born in North America between 2008 and 2016 and genotyped on the Axiom Equine Genotyping Array (Axiom MNEC670). All samples had a call rate >95%, and all SNPs had a call rate >95% and a minor allele frequency >5%. We called ROH with a minimum length of 300 kb using –homozyg in PLINK v1.90b (Purcell et al., 2007). All analyses were performed as previously described (Hill et al., 2022).

Inbreeding coefficients (F_{ROH}) were calculated by summing the total length of ROH for each individual and dividing by the autosomal genome length of 2281 Mb (McQuillan et al., 2008; Wade et al., 2009).

The distribution of F_{ROH} in North America was similar to the European and Australasian distribution, with median $F_{ROH} = 0.28$ (Figure S1). There was little evidence of population stratification in a principal component analysis using genotypes for European, Australasian, and North American samples, with the North American horses distributed within the variation for the other regions (Figure S2). On average, individual animals had 434 (range 370–670) ROH segments >300 kb with a mean length of 1.4 Mb (the longest ROH spanned 42 Mb) and mean F_{ROH} ranged from 0.16 to 0.34.

For horses born in North America and included in the study cohort, race records were obtained for Europe, Australasia, and North America. Among the genotyped horses, 654 had at least one recorded race start (raced) and 114 (14.8%) had no record of a race start (unraced). Raced horses had a median of 10 race starts. To test the hypothesis that inbreeding affects racing, we modelled the effects of inbreeding on racing using general linear

mixed models in lme4 (Bates et al., 2015). A horse was assigned 0 if it had never raced and 1 if it had raced. Among North American horses, F_{ROH} was not associated ($p = 0.518$) with the probability of ever racing (Figure S3, Table S1). Separating ROH into long (≥ 5 Mb) and short (<5 Mb) ROH also showed no effect on the probability of ever racing (Figure S4, Table S2).

In addition to modelling inbreeding effects on whether a horse had raced at all, we also fitted a model to test whether F_{ROH} is linked to a lower number of race starts among those horses that did race ($n = 654$). Among horses that did race, F_{ROH} was significantly associated ($p = 0.002$) with the number of race starts (Figures 1, S5, Table 1). The number of race starts decreased significantly ($p = 0.002$) with increasing F_{ROH} (log of the odds ratio: $\log_{OR}$ [95% confidence interval] $= -3.879$ [$-6.30, -1.46$], Table 1). Long ROH and short ROH had similar effects on the number of race starts (Figure S6, Table S3). Overall, horses with a 10% higher F_{ROH} than the mean inbreeding coefficient were predicted to have 3.5 fewer race starts compared to horses with a mean inbreeding coefficient (Table S4). The model predicts approximately 16 race starts for horses with the lowest F_{ROH} (0.18), but only seven starts for horses with the highest F_{ROH} (0.38). Approximately 9% of the total variance in number of races is explained by the model.

The results obtained for the North American population contrast with results for European and Australasian horses (Hill et al., 2022). Among European and Australasian horses, F_{ROH} was not associated with the number of race starts. For North America, inbreeding depression impacts on the durability (number of race starts) of horses that race, rather than on if they ever race at all. Furthermore, both long and short ROH contribute to the inbreeding depression observed for this trait in North America. The difference in trait effect may be considered in the context of regional variation in the racing ecosystems. One of the principal differences between racing in North America compared to Europe and Australia is the track surface on which horses train and race; in North America the majority of races are competed on dirt, whereas in Europe and Australia the majority of races are on turf (grass). Also, in North America, racing tends to be year-round at certain tracks at which horses are also stabled and trained. In Europe, horses tend to be stabled and trained in private yards away from the track and transported to race meetings that are held at particular times of year (Horse Racing Ireland, 2021), and in Australia there is a mixed model. Furthermore, there is also considerable variation in permitted and prohibited medication control across

FIGURE 1 Predicted counts of race starts for different inbreeding coefficients F_{ROH} using the model: races (>0)~F_{ROH}+sex+(1|birth year)+(1|olre). Poisson model for number of races among horses that raced. Olre is the observational level random effect account for overdispersion. Plots show the same prediction lines with (left) and without (right) raw data plotted on top. The right plot zooms closer into the relevant plotting area.
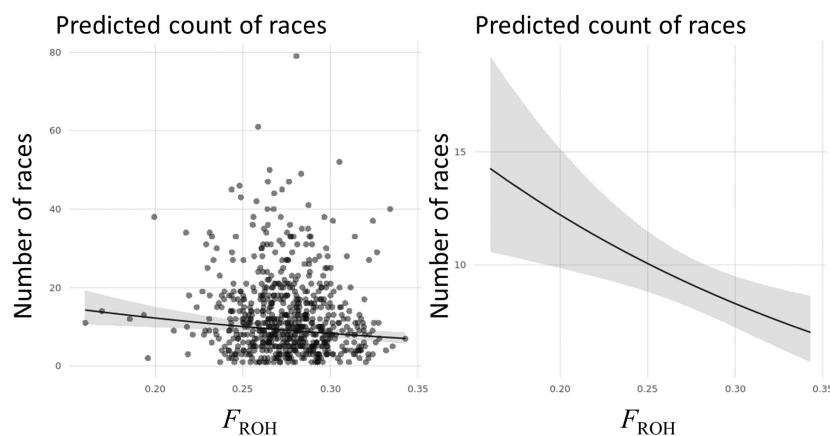

TABLE 1 Model estimates for a Poisson mixed model with the response variable 'race*s*' (*number of races*), for horses that had at least one race start using the model: races (>0)~Froh + sex + (1|birth_year)+(1|olre).

Term	Estimate $\log_{OR}$	CI (2.5%)	CI (97.5%)	z-value	p-value	Information
Intercept	3.279	2.608	3.95	9.573	1.70×10^{-17}	
Fixed effects						
F_{ROH}	−3.879	−6.296	−1.461	−3.144	0.0020	Continuous
Sex	0.387	0.246	0.528	5.385	7.24×10^{-08}	Categorical (0 = female, 1 = male)
Random effects (standard deviation)						
Birth year	0.141					$n = 9$
Observation-level random effect	0.677					$n = 768$

CI, confidence interval.

racing jurisdictions, both between countries and within the USA, across states (116th United States Congress [2019–2020], 2020).

A possible explanation for the association in North America between inbreeding and number of race starts, but not with ever racing, may be the practice of training and racing at the same race track. Horses that perform high intensity 'work' exercise training at the same race track at which they race could be more likely to start in a race than horses that are trained away from the race track and that need to be transported considerable distances (100s of km) to race. A second possibility is that our sample of North American Thoroughbreds is too small to reveal an effect on whether a horse ever raced, being only 13% of the sample size used for European and Australasian horses (Hill et al., 2022).

Irrespective of the explanation, here we show that inbreeding has a negative effect in the North American Thoroughbred, such that horses with higher levels of inbreeding are less durable than animals with lower levels of inbreeding. Considering the rising trend of inbreeding in the population (Binns et al., 2012; McGivney et al., 2020), these results indicate that there may also be a parallel trajectory towards breeding less robust animals. Indeed, over a 40-year period, the average number of race starts in North America has declined by 35% from 9.21 (1980) to 5.95 (2021) starts per runner (The Jockey Club, 2022). Although commercial considerations and opportunity to race may have influenced this decline, here we show that inbreeding is also a significant contributing factor, presumably as a result of the cumulative effect of deleterious mutations across the whole genome. Therefore, efforts to mitigate inbreeding should be considered for breed improvement. Managing inbreeding in the population by determining genetic relatedness among potential mate choices and selecting less related animals for breeding, could improve the racing durability of horses in North America with long-term positive effects on population health and sustainability.

KEYWORDS
inbreeding depression, North American Thoroughbred, race starts, runs of homozygosity

CONFLICT OF INTEREST STATEMENT
EWH is Chief Science Officer for Plusvital Ltd. BAM is an employee of Plusvital Ltd. EWH and DEM are shareholders in Plusvital Ltd. Other than the authors, the funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

DATA AVAILABILITY STATEMENT
The raw data analysed in this study is subject to the following licenses/restrictions: The phenotype and genotype data analysed in the present study are the property of Plusvital Ltd. and subject to a confidentiality agreement with the animal owners.

Emmeline W. Hill[1,2] 🔘
Beatrice A. McGivney[1]
David E. MacHugh[2,3]

[1]*Plusvital Ltd., The Highline, Dun Laoghaire Industrial Estate, Dublin, Ireland*
[2]*UCD School of Agriculture and Food Science, University College Dublin, Dublin, Ireland*
[3]*UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin, Ireland*

Correspondence
Emmeline W. Hill, Plusvital Ltd., The Highline, Dun Laoghaire Industrial Estate, Pottery Road, Dun Laoghaire, Dublin, A96 W5T3, Ireland.
Email: emmeline.hill@plusvital.com

ORCID
Emmeline W. Hill 🔘 https://orcid.org/0000-0002-1805-2250

REFERENCES
116th United States Congress [2019–2020]. (2020) H.R.1754 – Horseracing Integrity and Safety Act of 2020. https://www.congress.gov/bill/116th-congress/house-bill/1754/text

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.

Binns, M.M., Boehler, D.A., Bailey, E., Lear, T.L., Cardwell, J.M. & Lambert, D.H. (2012) Inbreeding in the thoroughbred horse. *Animal Genetics*, 43, 340–342.

Han, H., McGivney, B.A., Farries, G., Katz, L.M., MacHugh, D.E., Randhawa, I.A.S. et al. (2020) Selection in Australian thoroughbred horses acts on a locus associated with early two-year old speed. *PLoS One*, 15, e0227212.

Hill, E.W., Stoffel, M.A., McGivney, B.A., MacHugh, D.E. & Pemberton, J.M. (2022) Inbreeding depression and the probability of racing in the thoroughbred horse. *Proceedings: Biological Sciences, the Royal Society*, 289, 20220487.

Horse Racing Ireland. (2021) *Horse racing Ireland Factbook 2021.* Kildare, Ireland: The Curragh. https://www.hri.ie/hri/2021/HRI%20Factbook%202021%20FINAL.pdf

Lampi, S., Donner, J., Anderson, H. & Pohjoismaki, J. (2020) Variation in breeding practices and geographic isolation

drive subpopulation differentiation, contributing to the loss of genetic diversity within dog breed lineages. *Canine Med Genetics*, 7, 5.

McGivney, B.A., Han, H., Corduff, L.R., Katz, L.M., Tozaki, T., MacHugh, D.E. et al. (2020) Genomic inbreeding trends, influential sire lines and selection in the global thoroughbred horse population. *Scientific Reports*, 10, 466.

McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L. et al. (2008) Runs of homozygosity in European populations. *American Journal of Human Genetics*, 83, 359–372.

Orlando, L. & Librado, P. (2019) Origin and evolution of deleterious mutations in horses. *Genes (Basel)*, 10.

Petersen, J.L., Mickelson, J.R., Cothran, E.G., Andersson, L.S., Axelsson, J., Bailey, E. et al. (2013) Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PLoS One*, 8, e54997.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D. et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81, 559–575.

The Jockey Club. (2022) *2022 Fact book*. New York, NY, USA: The Jockey Club. https://www.jockeyclub.com/Default.asp?section=Resources&area=11

Todd, E.T., Ho, S.Y.W., Thomson, P.C., Ang, R.A., Velie, B.D. & Hamilton, N.A. (2018) Founder-specific inbreeding depression affects racing performance in thoroughbred horses. *Scientific Reports*, 8, 6167.

Wade, C.M., Giulotto, E., Sigurdsson, S., Zoli, M., Gnerre, S., Imsland, F. et al. (2009) Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*, 326, 865–867.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.